Winter 12-4-2017

Progress of Fintech Industry from Venture Capital Point of View

Ku-Hsien Hsieh
National Chengchi University, Taiwan, 103356506@nccu.edu.tw

Eldon Y. Li
National Chengchi University, Taiwan, eli@nccu.edu.tw

Hsieh, K.H. & Li, E.Y. (2017). Progress of Fintech industry from venture capital point of view. In *Proceedings of The 17th International Conference on Electronic Business* (pp. 297-301). ICEB, Dubai, UAE, December 4-8.

Progress of Fintech Industry from Venture Capital Point of View

(*Work in Progress*)

Ku-Hsien Hsieh*, National Chengchi University, Taiwan, 103356506@nccu.edu.tw
Eldon Y. Li, National Chengchi University, Taiwan, eli@nccu.edu.tw

ABSTRACT

Fintech (financial technology) is a term that broadly described the innovation of financial product and service using IT technology. Fintech uses a lot of new business model such as FXP2P and supply chain finance, and new technology such as blockchain and cryptocurrency. According to World Economic Forum report, fintech bring disruptive innovations that are reshaping the way financial services. Also, in World Economic Forum report have structured research framework against six (6) function of financial services and eleven (11) clusters of innovation. CrunchBase is a web 2.0 wiki-liked database for startup company that includes the information for founders, key team members, basic financial information, and venture capital funding and important events. This paper uses Python to web scrape CrunchBase website for studying the progress of Fintech industry and find out the development of IT technology and innovation. The objectives of this paper are to study the CrunchBase database to compare with World Economic Forum report research framework and find out the actual progress of Fintech industry.

Keywords: Fintech, Venture Capital, Innovation, Information Retrieval, Informetrics,

*Corresponding author

INTRODUCTION

Information science is a science field that study the collection, production, organization, classification, storage, retrieval, dissemination and protect of information. Information science can be a pure or as an applied science. Research in information science field can focus on the information user, the system that process information, the organization that adopt certain technology, or any other related field. Information Within information science, information retrieval and informetrics had developed from different areas, but share some common interest.

The informetrics traditionally can divided into six topics of: (1). Bibliometric theory, mathematical models and formalization of bibliometric laws, (2). Case studies and empirical papers, (3). Methodological papers including application, (4). Indicators engineering and data presentation, (5). Sociological approach to bibliometrics, sociology of science, and (6). Science policy, science management and general or technical discussion (Scharnhorst & Woulters, 2001). It is essential for informetrics that using mathematical or statistical model to analyze the information, and create indicators to present the usage of the information (Borner, Chen & Boyack, 2003).

Financial technology (Fintech) is a term that describe the new technology and innovation that aims to compete with traditional financial institutions to delivery better and more effective financial services. In today's digital era, people are seeking for easy access, convenience, efficiency, and speed. People want to use web or mobile to conduct transaction, and these transaction activities including the financial payment. Therefore, Fintech companies, mostly IT companies, start to provide both individuals and business with new financial tools or services. According to World Economic Forum report, fintech bring disruptive innovations that are reshaping the way financial services. Also, in World Economic Forum report have structured research framework against six (6) function of financial services and eleven (11) clusters of innovation (WEF, 2015; WEF, 2017).

As many new industries start to grow, the Venture Capital (VC) companies will start to invest in these start-up companies and help transfer into an industry. According to SparkLabs study, funding for VC-backed Fintech companies reaches to $13.8 billion in 2015 compared to $6.7 billion in 2014. The Fintech investment from VCs grows 106% from 2014 to 2015, just in one-year period. CrunchBase is a web 2.0 wiki-liked database for startup company that includes the information for founders, key team members, basic financial information, and venture capital funding and important events. This paper uses Python to web scrape CrunchBase website for studying the progress of Fintech industry and find out the development of IT technology and innovation (Srinaganya, 2015).

The objective of this paper is to design a model and indicator for Fintech using the concept and theory from Informetrics, then study the CrunchBase database to compare with World Economic Forum report research framework and find out the actual progress of Fintech industry.

The first section is the introduction. Then the paper is structured as the following: we will give brief background introduction in next section. Section 2 will include terminology of information retrieval (IR), informetrics/webometrics, fintech, and explain the investment progress of the venture capital industry. In Section 3, we will describe the World Economic Forum (WEF) research framework and the data structure of CrunchBase, then mapping the relationship between the date items. Then, the model design

and methodology are explained in section 4. Finally, discussion and implication are described in section 5.

TERMINOLOGY

Information Retrieval

Information retrieval (IR) is a selective process in which the information is extracted from a store information. The information can be structured, semi-structured, or non-structured. Traditionally, IR is used to locate the text-based information which can be any full-text documents or content type document such as bibliographic records of database. In recent years, the document had expanded to include websites HTML DOM-type documents and multimedia formats such as images, voices and videos (Ding, C, Foo, 2001)

The process of interactive information retrieval involves the searcher querying the IR system for the right information to be collected. The searcher initially submits a query or write a program to the IR system. The query search consists of one or more search terms that will be searched within the retrieved documents. A list of record matching the search query will then be stored into database for analysis. This process may repeat if necessary.

Recent years, IR system may also act as the search engines in World Wide Web (Almind & Ingwersen, 1997). In this case, the search data will be stated in HTML file text. Ever more, for web exchanged data format JSON, the search terms may be the JSON tag and search result will be the JSON data. Information retrieval remain a key research area within information science. Over the past twenty years, the information has expanded from simple text search, to DBMS search, then to World Wide Web search (Bennett & El-Arini, 2011).

Informetrics/Webometrics

Informetrics is quantitative study of information. Like information science, it includes the production, dissemination, and use of all form of information. Informetrics research investigates the existing data and attempts to develop mathematical models to better understand the information as whole. The informetrics encompasses several other fields: bibliometrics, scientometrics, webometrics, and cybermetrics (Thelwall, 2009; Hood, 2001). Bibliometrics studies the quantitative of literature records information. Scientometrics studies the statistical analysis of research pattern in sciences fields. Webometrics studies and measures the websites to get structure of the world wide web and usage patterns (Scharnhorst & Woulters, 2006).

Major areas of informetrics are (Wolfram, 2000):

1. Class bibliometric laws:
   - Lotka's Law – examines the author productivity in a frequency distribution
   - Bradfor's Law – examines the journal productivity as the concentration of articles in a subject area within journals
   - Zipf's Law – examines word usage frequency occurs within document
2. Citation and co-citation analysis – examines the citing patterns of authors and publications to determine the importance of the paper
3. Indicators – examines the productivity of studies area output
4. Information growth and obsolescence – examines subject area grows or decline over the time
5. Information resource usage – examines the information resources are used over time

Venture Capital

Venture Capital is a type of private equity to provide the funding for the startup companies, possible in highly growth potential industries. Investing in early-stage companies is risky. The venture capital investors are typically very large financial institutions such as pension fund, insurance companies, and private bank. The VCs put small percentage of their investment into high-risk start-up companies. The purpose of VCs investment is hoping to get sufficient returns at acceptable risk. According to Harvard Business Review articles, the VCs in general expect a return of 25~35% per year over the lifetime of the investment (Zider, 1998).

The answers for VCs to meet the expected return are controlling the investment profile and the deal structure. The priority for the investment profile is to invest in good industry, not good people and good idea. So, it is essential for VCs to invest the companies in accelerating growth period.

The is crucial to understand how venture capital performs the "funding rounds" progress. In general, the funding rounds will go through the different periods in the following table. The funding grows rapidly when funding round progresses. In general, the seed round is under US$500K, Series A then will move above US$ 1 million. The venture capitalist can exit during Mezzanine round, IPO, or an acquisition (Zider, 1998).

Table 1: Venture capital funding process

Funding Period	Sources
Seed Round	Family and Friends, Angel investors, Accelerators, incubators
Series A (1st Round)	Venture capital (VC)
Series B (2nd Round)	Venture capital (VC)
Series C (3rd Round)	Venture capital (VC)
Mezzanine Round	Venture capital (VC)

Fintech Innovation

According to Ireland National Digital Research Centre had defined Fintech as a "Financial Service Innovation". Fintech can be viewed as a new type of solution innovation. This innovation can be any form of financial service such as business model, financial instrument, process, application development, or any other financial related services. These Fintech innovations develop new attraction to the customers and disrupt the traditional financial companies (Li & Li, 2016). From Accenture's "Fintech Innovation labs eligibility criteria" report had stated the global investment in Fintech had grown from $9.3 billion in 2008 to $120 billion in 2015, 12 times of growth.

RESEARCH DESIGN

WEF Innovation Research Framework

Because Fintech industry had growing so fast, many studies had done to emphasize the innovation within the financial services industry. The most popular Fintech Innovation framework is World Economic Forum's 2015 report – "The Future of Financial Services". WEF report identified 11 key clusters of innovations based the impact to six core functions of financial services as the following figure 1.

Figure 1: WEF Fintech innovation research framework (WEF, 2015)

CrunchBase

CrunchBase is a database of the startup companies that include the company profile, the founder, the management team, the funding rounds and the investors. The CrunchBase data items is listing the following table 2.

Table 2: CrunchBase data items

Date Section	Date Items
Company Profile	● Company name ● Company location ● Industry category ● Tag keyword ● Founded date ● Operating status ● Competitor
Funding Rounds	● Number of funding rounds ● Total funding amount ● Funding date – Funding amount ● Funding investors
Investors	● Number of lead investors ● Number of investors
Investor Profiles	● Investor name ● Investor categories ● Investor founded date ● Investor type

Model Design

By categorized the Fintech innovation areas and categories from WEF report research framework, we can use CrunchBase database to form our research model. The model design is shown in the following figure. The research design model is following the informetrics research study model. In this paper, we use Python computer language as the research model process language. First, the data is collected. In this stage, the data is collected by using information retrieval web crawling technique, with Python crawling package such as Scrapy and BeautifulSoup. Second, after the data is loaded into Python. We will map the correlation between informetrics data item to CrunchBase data items. In the data item mapping, the startup company is like the article, the company industry categories are like journal name, and the company service category is like the keyword of the article. So, we can easily use informetrics model to process and validate the data. Third, we will add the venture capital profile into the model. Finally, we build the indicator for the startup companies against the venture capital. It is much like citation. We can easily to find which companies got more funding, or as in informetrics term, more citation. Then, we can add the impact factor label for these Fintech companies.

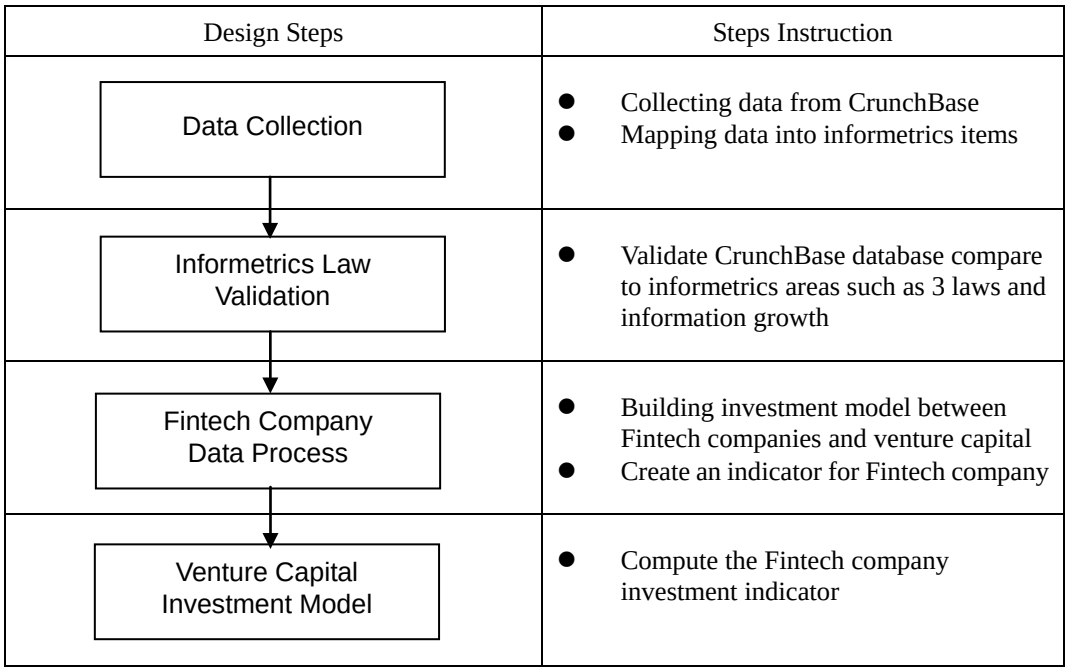

Figure 2: Fintech company indicator research model

CONCLUSION

At the beginning of the conclusion, the author has to say this paper is still a working paper. The model is under development and current progress is presented as the following. First, the author had using Python to parse the company profile that label "Fintech". Total of 4,730 Fintech companies had founded from CrunchBase database and totally of 154 tag keywords were indexed, and more than $200 billion was funded. Second, the data from Crunchbase basically followed the basic principle laws of informetrics. Of course, still more research needs for developing the indicator and more computation needs for validate the venture capital investment model.

REFERENCES

[1] Almind, T. & Ingwersen, P. (1997) 'Informetric analyses on the world wide web: methodological approaches to 'webometrics'', *Journal of Documentation*, Vol. 53, No.4, pp. 404-426.

[2] Bennett, P. & El-Arini, K. (2011) 'Enriching Information Retrieval', *ACM SIGIR Forum*, Vol. 4, No. 2, pp.60-64.

[3] Borner, K., Chen, C. & Boyack, K.W. (2003) 'Visualizing Knowledge Domains', *Annual Review of Information and Technology*, Vol. 37, No. 1, pp. 179-255.

[4] Ding, Y., Chowdhury, G.G. & Foo S. (2001) 'Bibliometrics Cartography of Information Retrieval Research by using Co-word Analysis', *Information Processing & Management*, Vol. 37, No. 6, pp. 817-842.

[5] Hood, W. W., & Wilson, C. S. (2001) 'The Literature of Bibliometrics, Scientometrics, and Informetrics', *Scientometrics*, Vol. 52, No. 2, pp. 291-314.

[6] Li, E. & Li, J. 'Information Security Challenges in the New Era of Fintech', *Proceeding of ICEB 2016*.

[7] Scharnhorst, A. & Woulters, P. (2006) 'Web indicators – A New Generation of S&T Indicators?', *Cybermetrics*, Vol. 10, Issue: 1.

[8] Srinaganya. G. & J.G.R. S. (2015) 'A Technical Study on Information Retrieval using Web Mining Techniques', *ICIIECS, IEEE*.

[9] Thelwall, M. (2009) *Introduction to Webometrics: Quantitative Web Research for the Social Sciences*, San Raphael, CA: Morgan & Clapypool.

[10] World Economic Forum (2015) *The Future of Financial Services*, World Economic Forum.

[11] World Economic Foru. (2017) *Beyond Fintech: A Pragmatic Assessment of Disruptive Potential In Financial Services*, World Economic Forum.

[12] Wolfram, D. (2000) 'Application of Informetrics to Information Retrieval Research', *Informing Science*, Vol. 3, No. 2, pp. 77-82.

[13] Zider, B. (1998) 'How Venture Capital Works', *Harvard Business Review*, Issue: November-December, 1998.